

Relativizing an incompressible number and an incompressible function through subrecursive extensions of Turing machines[☆]

Felipe S. Abrahão^a

^a*Postdoctoral researcher at the National Laboratory for Scientific Computing (LNCC), Brazil*

Abstract

We show in this article that uncomputability is also a relative property of subrecursive classes built on a recursive relative incompressible function, which acts as a higher-order “yardstick” of irreducible information for the respective subrecursive class. We define the concept of a Turing submachine, and a recursive relative version for the Busy Beaver function and for the halting probability (or Chaitin’s constant) Ω ; respectively the Busy Beaver Plus (BBP) function and a time-bounded halting probability. Therefore, we prove that the computable BBP function defined on any Turing submachine is neither computable nor compressible by any program running on this submachine. In addition, we build a Turing submachine that can use lower approximations to its own time-bounded halting probability to calculate the values of its Busy Beaver Plus function, in the “same” manner that universal Turing machines use approximations to Ω to calculate Busy Beaver values. Thus, the algorithmic information carried by the BBP function is relatively incompressible (and uncomputable) at the same time that it still is occasionally reached by submachines. We point that this phenomenon enriches the research on the relativization and simulation of uncomputability and irreducible information.

Keywords: Algorithmic information, Subrecursion, Relative computability, Busy Beaver function, Halting probability, Time-bounded Turing machines
2010 MSC: 68Q30, 68Q05, 03D20

1. Introduction

Let’s remark some important concepts and definitions introduced in [1, 2]. Basically, we can define a **Turing submachine** as a Turing machine that always gives an output for any input, i.e. always halts. So, note that Turing submachine

[☆]This research was supported by the National Council for Scientific and Technological Development (CNPq), Brazil, as a PhD Fellowship at the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil.

is just another terminology for **total Turing machines** [6, 22]. However, despite the fact that they would be just different names for the same object and can be used interchangeably, the expression “total Turing machine” might not immediately capture its relevant properties related to the present work.

Remember that every computable total function (or total Turing machine) defines a subrecursive class which is a proper subclass of others subrecursive classes (and of the class of all recursive functions) [5, 9, 23, 28]. A subrecursive class is one defined by a proper subset of the set of all problems with Turing degree **0**. Therefore, a **subcomputable** class of problems will be subrecursive, because it will never contain all recursive/computable problems. The term subrecursion is also used to characterize subrecursive hierarchies, as in Kleene and Grzegorzczuk [5, 20, 27], covering all primitive recursive functions. But for us, the prefix refers more specifically to the concept of subrecursive class [5, 20, 21].

For example, the total Turing machine may be a subsystem of another total machine which is capable of computing functions that are relatively uncomputable by the former. This very idea of being part of another non-reducibly more powerful machine is the core notion of the expression “submachine”, conveying and bearing the ideas of hierarchies of subrecursive classes together with the powerful concept of Turing machines. Thus, the terminology Turing submachine emphasizes this property of total Turing machines being always able to be part of another proper and bigger machine. For more of this discussion, see [2] and section 4.

The main idea presented here is building, and rather proving, a system (a Turing machine) that can “behave” in relation to a subsystem (its Turing submachine) in the same way as a hypercomputer [15, 29] (or an oracle Turing machine [30]) would behave in relation to a subsystem (in particular, a universal Turing machine). In fact, we will not emulate all — which might be impossible [20, 21] — the properties of a hypercomputer in relation to a computer, but focus on defining a function $BB^{+}_{P^{**}_{T \circ P_T}}(N)$ analogous to a Busy Beaver function (in Chaitin’s first models [10] in metabiology [11, 12], the function $BB(N)$), and a recursive relative number $\Omega_{P^{**}_{T \circ P_T}}$ analogous to Ω , so that this function will behave in relation to the Turing submachine and to the number $\Omega_{P^{**}_{T \circ P_T}}$ “resembling” the way the original Busy Beaver behaves in relation to a universal Turing machine and to Ω .

In short, $BB^{+}_{P^{**}_{T \circ P_T}}(N)$ must be relatively uncomputable and incompressible by any **subprogram** (a program running on a Turing submachine), the same — except for a constant — way the original Busy Beaver is incompressible by any program running on a universal Turing machine. We will call this phenomenon of relativization of the uncomputability of a hypercomputable function defined on Turing submachines as **recursive relative uncomputability**, or **sub-uncomputability**, of a function (in our case, Busy Beaver function).

Also bear in mind that for the present purposes we do not want it to be “overly” relatively uncomputable, as a first order oracle machine is already capable of computing $BB'(N)$. We want the new relative Busy Beaver function to be “as, but not more than”, uncomputable for a Turing submachine as is the

Busy Beaver for a universal Turing machine. This way, it keeps playing a role of an outside measurement of irreducible complexity (or information) [10, 32] while being occasionally reached by some subprograms. For this purpose, we will demonstrate that there are subprograms that calculate $BB^{+}_{P^{**}_T \circ P_T}(N)$ values for lower approximations to $\Omega_{P^{**}_T \circ P_T}$, as is already known in the case of the Busy Beaver and of Ω .

Before reaching the three theorems that this article aims, we start by defining some general conditions on a self-delimiting programming language for the universal Turing machine U . The definition of the Turing submachine $U_{P^{**}_T \circ P_T}$ playing the central role herein is intrinsically also dependent on the time-bounded halting probability $\Omega_{P^{**}_T \circ P_T}$ and on the function $BB^{+}_{P^{**}_T \circ P_T}$, since the algorithm of the program $P^{**}_T \circ P_T$ is defined by a successive and self-referential composition of other algorithms — a construction that resembles a fixed-point (or diagonal) lemma [17, 26] used to prove incompleteness results for example. Then, we prove that $U_{P^{**}_T \circ P_T}$ is indeed well-defined. It relies on the fact that bigger subprograms only depends on smaller subprograms to always halt, which enable us to build a proof by induction. The second theorem derives from the very construction of the program $P^{**}_T \circ P_T$ and from the former result. It proves the desired property of, for every N , function $BB^{+}_{P^{**}_T \circ P_T}$ being calculated by some subprograms through lower approximations to $\Omega_{P^{**}_T \circ P_T}$ as inputs. And we finish by proving the uncomputability and incompressibility of the function Busy Beaver Plus by any program running on a Turing submachine. This proof, as previously presented in [1, 2], is based on an argument analogous to the one in Chaitin’s incompleteness theorem [13], which uses algorithmic complexity/information. Therefore, we conclude by showing that function $BB^{+}_{P^{**}_T \circ P_T}$ is sub-uncomputable by any program running on $U_{P^{**}_T \circ P_T}$, but still reachable by giving lower approximations to $\Omega_{P^{**}_T \circ P_T}$ as inputs to a class of programs running on $U_{P^{**}_T \circ P_T}$.

2. Language L

It is important to us that the submachine $U_{P^{**}_T \circ P_T}$ can be programmable. Its language must be able to be used on any usual computer. Hence, the properties and rules of well-formation of the universal programming language for the chosen universal Turing machine U upon which we will define submachines must be recursive. As one is invited to see below in definition 2.1, the funding conditions for the language L are actually met by most self-delimiting programming languages for practical purpose. As we already have shown in [2], this leads us, along with theorem 11.1, to the conclusion that the phenomenon of sub-uncomputability is ubiquitous.

Since we are trying to build a computer that can emulate some “behaviors” of a first order oracle Turing machine with respect to a universal Turing machine, it is necessary that one can “teach” the machine to perform and recognize “well-formed functionalizing concatenations” within the language. That is, one needs a language that provides a direct way of symbolizing a program taking any given bit string as input (for example, a program p that is, actually, program p' taking

program p' as input), which makes this program act as a function. Note that this type of program is already used to demonstrate the halting problem, or to demonstrate that the Busy Beaver function is uncomputable for example. But the form it may assume is completely arbitrary, as a universal Turing machine — note that it is not our case for the present purposes —, in any case, will run it.

The desired condition is met by a recursively well-formed type of concatenation, which we will symbolize by “ \circ ”. And we denote the optimal functionalizing concatenation, which is “joining” strings in the most compressed way possible, by “ $*$ ”.

2.1. Definition

We say a universal programming language l , defined on a universal Turing machine U , is **recursively functionalizable** if there is a program that, given any bit strings P and w as inputs, returns a bit string belonging to l which will be denoted as $P \circ w$, whereby $U(P \circ w)$ equals “the result of the computation (on U) of program P when w is given as input”. In addition, there must be a program that determines whether or not a bit string is in form $P \circ w$ for any P and w , and is capable of returning P and w separately.

Analogously, it also must be true for the successive concatenation $P \circ w_1 \circ \dots \circ w_k$, with program P receiving w_1, \dots, w_k as inputs.

So, the general definition of our language L comes just below.

Let U be a universal Turing machine running language L , a universal language that is **binary**, **self-delimiting**, **recursive**, and **recursively functionalizable** such that there are constants $\epsilon, C \in \mathbb{C}'$, for every P, w_1, \dots, w_k , where:

$$|w_i| < |P \circ w_1 \circ \dots \circ w_k|, \text{ for } i = 1, 2, \dots, \text{or } k$$

and

$$|P \circ w_1 \circ \dots \circ w_k| \leq C \times k + |P| + |w_1| + |w_2| + \dots + |w_k|$$

and

$$H(N) \leq C' + \log_2 N + (1 + \epsilon) \log_2 (\log_2 N)$$

Let $H(w)$ denote the algorithmic complexity (or Solomonoff-Kolmogorov-Chaitin complexity) of the bit string w defined on the machine U , i.e. $H(w)$ is the size of the shortest program P in L such that $U(P) = w$ (see definition 3.b). And let $H(N)$ denote the algorithmic complexity of the natural number N , that is, the size of the shortest program in L that outputs¹ N when run on the machine U — see definition 3.d.

¹A representation (except for a trivial bijection) of N in the language L actually.

3. Some definitions and notations

- a) W is the set of all finite bit strings, where the computable enumeration of these bit strings has the form $l_1, l_2, l_3, \dots, l_k, \dots$
For practical purposes, a language may be adopted where $l_1 = 0$.
- b) Let l be a language for a universal Turing machine U . Let $w \in l$.
Then, $U(w)$ denotes the result of the machine U running the string w .
- c) Let $w \in W$.
Then, $|w|$ denotes the size or number of bits contained in w .
- d) Let N simply symbolize the corresponding program in language L for the natural number N . For example, $P \circ N$ denotes program $P \circ w$ where w is a representation of the natural number N in the language L .
- e) If function f is computable by program P , then f may also be called function P .

4. Turing submachines

As discussed in [2], we follow the conventional understanding in which a computation that is a part of another computation may be called a subcomputation, and a machine that is a part of another machine may be called a submachine. For example, a Turing submachine can be a program or subroutine that the “bigger” Turing machine runs, always generating an output, while performing various other tasks. Note that it is true (a theorem) that for every total Turing machine there is another Turing machine that completely simulates and contains the former total Turing machine, in a manner that the computations of the latter contains the computations of the former. So, a system can be taken as a Turing machine, and a subsystem can be taken as a Turing submachine.

In fact, we are using a stronger notion of subsystem based upon this conventional notion: a subsystem must be only able to do what the system knows, determines and delimits. This way, submachines will only be those machines for which there is another non-reducibly “bigger” machine that can decide, at least, what is the output of the former and whether there is an output at all. In particular, this condition proved to be necessary in order to build theoretical models for an open-ended² [18] evolution of programs in which the very environment, or “Nature”, can be simulated in a computer [1]. Note that every machine that falls under this definition always defines an equivalent **total Turing machine** (with a signed output corresponding to the case where the submachine does not halt); and every total Turing machine falls under this definition.

Let P_f be a program running on U defined in language L , computing a total function (a function defined for all possible input values) f such that $f : L \rightarrow X \subseteq W$. The language W does not need necessarily to be self-delimiting, and may be comprised of all bit strings of finite size, as long as they

²That is, an evolution of software through random algorithmic mutations in which programs (“organisms”) always grow in algorithmic complexity (or gain more irreducible information) over time.

may be recursively enumerated in order, as l_1, l_2, l_3, \dots . For practical reasons, we will choose an enumeration where $l_1 = 0$.

A “**Turing submachine**”, or **total Turing machine**, U/f is defined as a Turing machine in which, for every bit string w in the language of U , $U/f(w) = U(P_f \circ w)$.

This definition is quite general and transforms any total computable function into a Turing submachine. In fact, as said in the introduction, Turing submachines are just another name for total Turing machines. Anyway, submachines can always be subsystems of either abstract universal Turing machines or of powerful enough everyday computers (which are also some sort of total Turing machine, i.e. a universal Turing machine with limited resources).

Note that the class of all submachines is infinite, but not recursive. And when we talk of **subprograms** we refer to programs run on a Turing submachine.

We will now use another concept of vital importance: **computation time**. Similarly to time complexity, we will call T a program that calculates how many steps or basic operations U performs when running program p . Thus, if $U(p)$ does not halt, then $U(T * p)$ will not halt either, and vice versa. Herein, only submachines of a particular subclass will be dealt with: submachines defined by a computation time limited by a computable function, i.e. time-bounded Turing machines [19]. In fact, both these and the more generic submachines defined above are equivalent in computational power — although they might not be in definition. To demonstrate this, just note that if a program computes a total function, then there is a program that can compute the computation time of this first program. Therefore, for every computable total function, there is a submachine with bounded computation time capable of computing this function (but possibly other functions as well). The reverse directly follows from the definition of submachine. This holds even though the definition of a submachine may come from more arbitrary restrictive conditions (although still recursive) than just limiting the computation time by some upper bound gave by a computable function (e.g. a polynomial one).

Let P_T be an arbitrary program that calculates a computation time for a given program w . That is, let P_T be an arbitrary computable total function. Thus, there is a **Turing submachine U_{P_T} defined by the computation-time function P_T** .

We then define submachine $U/P_{SM \circ P_T}$ (which will be a program running on U that computes a total function), where P_{SM} is a program that receives P_T and w as inputs, runs $U(P_T \circ w)$ and returns:

- (i) l_1 , if $U(w)$ does not halt within computation time $\leq U(P_T \circ w)$;
- (ii) l_{k+1} , if $U(w)$ halts within computation time $\leq U(P_T \circ w)$ e $U(w) = l_k$;

This program defines a Turing submachine that returns a known symbol (in our case, from definition 3.a, we have that $l_1 = 0$) when program w does not halt in time $\leq U(P_T \circ w)$ or returns the same output (except for a trivial bijection) as $U(w)$ when the latter halts in time $\leq U(P_T \circ w)$.

To be a Turing submachine, $U/P_{SM \circ P_T}$ must be defined for all inputs. This occurs because P_T is total by definition. In addition, as computation time

P_T becomes more increasing, the more submachine $U/P_{SM} \circ P_T$ approaches the universality of U .

Therefore, we will denote only as U_{P_T} a Turing submachine $U/P_{SM} \circ P_T$, so that:

$$\forall w \in L \ (\ U_{P_T}(w) = U / P_{SM} \circ P_T (w) = U (P_{SM} \circ P_T \circ w))$$

5. Function $BB^+_{P'_T}$

Let P'_T be a total function. Let us define function $BB^+_{P'_T}(N)$, which we call **Busy Beaver Plus**, through the following recursive procedure:

- (i) Generate a list of all outputs of $U_{P'_T}(w)$ such that $|w| \leq N$;
- (ii) Take the largest number on that list;
- (iii) Add 1;
- (iv) Return that value.

The name of this function refers to the Busy Beaver $BB(N)$ function and, consequently, it is no coincidence that the two have almost the same definition. If step (iii) is removed, it becomes exactly the Busy Beaver function for Turing submachines (denoted below as $BB_{P'_T}(N)$). Thus:

$$BB^+(N) = BB(N) + 1$$

and

$$BB^+_{P'_T}(N) = BB_{P'_T}(N) + 1$$

But why use function BB^+ instead of BB ? As we are dealing with Turing submachines and P'_T is arbitrary, it is possible that there is a program on $U_{P'_T}$ with size $\leq N$ such that computes the highest value returned by any other program on $U_{P'_T}$ with size $\leq N$. When dealing with a universal Turing machine U , this cannot occur (except for a constant). However, with submachines, it can. Thus, function $BB^+_{P'_T}$ is triggered to assure it, in itself, is not relatively computable (and compressible) by any program on $U_{P'_T}$, although it can be by a program on U . Note that, since P'_T is a program that computes a total function, then $BB^+_{P'_T}(N)$ is computable.

The Busy Beaver contains the idea of the greatest output of any $\leq N$ sized program; so the Busy Beaver Plus function contains the idea of surpassing (by 1) any $\leq N$ sized program. Respectively, the first gives us maximization, and the second, an “almost” minimal increment. No matter how rapidly increasing is the function P'_T , the program on U that computes $BB^+_{P'_T}(N)$ simply bases itself on the $U_{P'_T}$ outputs to overcome them by a minimum. No matter how powerful $U_{P'_T}$ may be, $BB^+_{P'_T}(N)$ will always be “one step” ahead of the best that any subprogram (i.e., any program $U_{P'_T}$) can do.

It is worthy of note that, analogously to the Busy Beaver, the $BB^+_{P'_T}(N)$ may be used to measure the algorithmic creativity or **sub-algorithmic complexity** [20] of the subprograms in relation to Turing submachine $U_{P'_T}$, like a

resource-bounded algorithmic complexity [3, 24, 31]. Why? By its very definition, if a subprogram generates an output $\geq BB^+_{P'_T}(N)$, it must necessarily be of size $> N$. It needs to have more than N bits of relatively incompressible information, i.e. more than N bits of relative creativity. Further along, we will build a submachine $U_{P^{**}_T \circ P_T}$ that always runs a subprogram of size $\leq 2N + C$ that calculates $BB^+_{P^{**}_T \circ P_T}(N)$, where C is a constant. Therefore, the sub-algorithmic complexity $H_{P^{**}_T \circ P_T}$ of function $BB^+_{P^{**}_T \circ P_T}(N)$ is such that

$$N < H_{P^{**}_T \circ P_T}(BB^+_{P^{**}_T \circ P_T}(N)) \leq 2N + C$$

In summary, as we will prove below in theorem 10.2, this allows submachine $U_{P^{**}_T \circ P_T}$ to occasionally grasp irreducible information “from the outside” through lower approximations to $\Omega_{P^{**}_T \circ P_T}$.

Of course, one may always build a program that computes function $BB^+_{P'_T}$, if function P'_T is computable. This would allow a far smaller program than N there to exist — e.g., of size $\leq C + \log_2 N + (1 + \epsilon)\log_2(\log_2 N)$ — that will compute $BB^+_{P'_T}(N)$. But that does not constitute a contradiction, because this program can never be a subprogram of $U_{P'_T}$, in other words, it can never be a program that runs on the computation time determined by P'_T . If it was, it would enter into direct contradiction with the definition of $BB^+_{P'_T}$: the program P'_T will become undefined for an input, which by assumption is false.

Further ahead the reader may attest that this function is the same as the one defined by $U(\pi'_\Omega \circ P'_T \circ 0^{U(P'_\Sigma \circ P'_T \circ N)})|_1 \circ U(P'_\Sigma \circ P'_T \circ N)$. This is deliberate, as it makes π'_Ω and $BB^+_{P'_T}(N)$ resemble π_Ω and $BB(N)$ respectively, as introduced in [1]. This is what allows us to simulate some hypercomputable properties of a first order hypercomputer in relation to a computer.

6. Program P_Σ

Let P_Σ be a program that takes $P'_T \circ N$ as input and calculates the binary sum of all algorithmic probabilities of the programs with size $\leq N$ in the language L that halt in computation time defined by P'_T (or, $U_{P'_T}(w) \neq l_1$). If $N = 0$, the output is 0.

If P'_T is not a total function (i.e., does not halt for an input w), then P_Σ will also not be defined for all inputs. If P'_T is, then obviously P_Σ will also be.

Remember, this is a self-delimiting language; therefore, the output of P_Σ will always be a real binary number between 0 and 1.

But what happens if we increase the number N progressively? As the sum of the algorithmic probabilities of all L programs is ≤ 1 and $U(P_\Sigma \circ P'_T \circ N)$ is always non-decreasing, will there be a limit in $U(P_\Sigma \circ P'_T \circ N)$ when N tends to infinity (as long as P'_T is a total function). What does this limit reveal? It will be, precisely, the time-bounded version of Chaitin’s Omega number.

7. A relative Chaitin's constant: the time-bounded halting probability

Through the construction of program P_{SM} , it is possible to note that, in the limit, P_Σ calculates the sum of all programs in L that halt in computation time determined by P'_T . With this, a direct analogy with Chaitin's Omega number [7, 8, 14] is obtained, which justifies calling this number a time-bounded halting probability.

Let P'_T be a total function. So $\Omega_{P'_T}$ may denote the **time-bounded halting probability** defined by:

$$\Omega_{P'_T} = \sum_{p \text{ is a program in } L \text{ that halts in computation time } \leq U(P'_T \circ p)} 2^{-|p|}$$

In other words, p is counted in this sum if, and only if, for every y and x ,

$$U(T \circ p) = x \wedge U(P'_T \circ p) = y \rightarrow x \leq y$$

Thus, the more rapidly increasing is the function given by program P'_T , the closer Ω will be approached. In addition, note that non-halting programs p are excluded: after y computations, if program p has not halted yet, it can only either never halt, or halt in time x , with $x > y$.

8. Subprograms $\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho$

Program $\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho$ will now be explained. What does this program do? It is almost the same program π_Ω used by Chaitin in his early models for cumulative evolution in metabiology [10]. However, we take here as reference a Turing submachine $U_{P'_T}$ instead of the universal machine U . Furthermore, the output is not approximation indices to $\Omega_{P'_T}$, but approximations to $BB^+_{P'_T}$. This is equivalent if one is dealing with the classical halting probability Ω or the classical Busy Beaver function $BB(N)$, which are defined on U . However, it does not necessarily holds if one is dealing with Turing submachines.

Let P'_T be a total function.

Program π'_Ω reads bit string $P'_T \circ 0^{|\rho|} 1 \circ \rho$ as its input and sums all the algorithmic probabilities of programs of size $\leq n$ that have halted in the computation time determined by P'_T . It starts with $n = 1$ and continues up to k so that the sum of all the algorithmic probabilities of all programs of size $\leq k$ is equal to or larger than the value of ρ . Then it calculates the largest output among these subprograms of size $\leq k$, adds 1 and returns this value. If $\rho = 0$, then it returns 0.

Note that if $\rho > \Omega_{P'_T}$, this program will never halt. In case $\rho = \Omega_{P'_T}$ — which would be an ideal possibility — π'_Ω may, or may not, halt: it will only not ever halt if $|\rho| = |\Omega_{P'_T}|$ is infinite, since π'_Ω would never finish reading its input when $\Omega_{P'_T}$ is of infinite size. If ρ has finite size and $\rho \leq \Omega_{P'_T}$, then π'_Ω halts.

Hence it may be concluded that π'_Ω takes a finite lower approximation to $\Omega_{P'_T}$ and calculates a value that exceeds the value of any output produced, by any program involved, in ascending order of size, over the sum of algorithmic probabilities that results in the lower approximation in question. This is practically analogous to what happens with π_Ω in relation to Ω and $BB'(N)$ in Chaitin's work [11], except for a increment (which is adding 1).

Thus, assuming P'_T is total, it can be shown that $\pi'_\Omega \circ P'_T \circ 0^{|\rho'|} 1 \circ \rho'$ always halts when $\rho' \leq U(P_\Sigma \circ P'_T \circ N)$, for every N . And it does not matter if $|\rho'| > |U(P_\Sigma \circ P'_T \circ N)|$, as what is being dealt with are real numbers between 0 and 1. For example, the trailing zeros do not affect the fact that $\rho' \leq U(P_\Sigma \circ P'_T \circ N)$ — and this will be important later on.

9. Submachine $U_{P^*_T \circ P'_T \circ P_T}$

Now we will define a Turing submachine $U_{P^*_T \circ P'_T \circ P_T}$. Note that P'_T and P_T are two total functions, by assumption. This will remain until we reach submachine $U_{P^*_T \circ P_T}$, where P_T is still total, by assumption, but in place of P'_T , it will be self-referentially put $P^*_T \circ P_T$. Also remember that program T receives w as input and returns the computation time of $U(w)$.

Program P^*_T receives P'_T , P_T and w as inputs, and returns:

- (i) $U(T \circ w)$, if w is in the form $\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho$ and

$$U(P_\Sigma \circ P'_T \circ (|\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) \geq \rho$$

and

$$U(T \circ w) > U(P_T \circ w);$$

- (ii) $U(P_T \circ w)$, if case (i) does not apply;

Here a program is being defined upon two total functions: P'_T and P_T . The first step to verify whether or not P^*_T is total is noting that there is a program that always decides if w lies in case (i) or (ii).

The program must first verify the form of the program w . For case (i), if

$$U(P_\Sigma \circ P'_T \circ (|\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) \geq \rho$$

we will have that $U(T \circ \pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho)$ is always well defined. The only thing left to determine is whether

$$U(T \circ \pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho) > U(P_T \circ \pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho)$$

or not, which is also decidable. On the other hand, if

$$U(P_\Sigma \circ P'_T \circ (|\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) < \rho$$

then case (ii) applies directly. Case (ii) derives directly from the fact that P_T is also total.

It is important to attend to the fact that $P^*_T \circ P'_T \circ P_T$ is sufficient computation time to compute anything that U_{P_T} may compute. In case (ii) this is obvious. In case of an (i)-formed program, condition $U(T \circ w) > U(P_T \circ w)$ will always be present ensuring that $P^*_T \circ P'_T \circ P_T$ is always an extension of P_T , and never a restriction.

Summarizing what $P^*_T \circ P'_T \circ P_T$ does in case (i), $U_{P^*_T \circ P'_T \circ P_T}$ always allows, at least, a program of the form $\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho$, which will calculate a value always higher than any output from any program on $U_{P'_T}$ of size $\leq N$, so that $U(P_\Sigma \circ P'_T \circ (|\pi'_\Omega \circ P'_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) \geq U(P_\Sigma \circ P'_T \circ (N))$. As we will prove, this essentially always enables submachine $U_{P^*_T \circ P'_T \circ P_T}$ to calculate $BB^+_{P'_T}(N)$ with a subprogram of size $\geq N + 1$ and $\leq 2N + C$, where C is a constant.

Of course, once that for every w , $U(P_T \circ w) \leq U(P^*_T \circ P'_T \circ P_T \circ w)$, then

$$\Omega_{P^*_T \circ P'_T \circ P_T} \geq \Omega_{P_T}$$

However, what about if our objective is for $U(P_\Sigma \circ P'_T \circ N)$ to always be an approximation to $\Omega_{P^*_T \circ P'_T \circ P_T}$ itself, rather than an approximation to $\Omega_{P'_T}$? That is, in case (i) that $U_{P^*_T \circ P'_T \circ P_T}$ always be able to calculate values of $BB^+_{P^*_T \circ P'_T \circ P_T}$ with a subprogram of size $\geq N+1$? Which program should replace for P'_T ? Here the employment of a self-diagonalizing self-reference will become evident, within program P^*_T itself, to build program P^{**}_T below.

10. Submachine $U_{P^{**}_T \circ P_T}$

Let P_T be a total function. Program P^{**}_T receives P_T and w as inputs, reads itself, P_T and w , and then assembles $P^*_T \circ P^{**}_T \circ P_T \circ P_T \circ w$ and returns $U(P^*_T \circ P^{**}_T \circ P_T \circ P_T \circ w)$.

That is, $U(P^{**}_T \circ P_T \circ w) = U(P^*_T \circ P^{**}_T \circ P_T \circ P_T \circ w)$, returning:

- (i) $U(T \circ w)$, if w is in form $\pi_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$ and

$$U\left(P_\Sigma \circ P^{**}_T \circ P_T \circ (|\pi_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)\right) \geq \rho$$

and

$$U(T \circ w) > U(P_T \circ w);$$

- (ii) $U(P_T \circ w)$, if case (i) does not apply;

The issue at hand is whether this program is a total function or not. It will not be as simple as in the previous case. Case (ii) remains trivial, since it depends only on P_T . In (i), the problem is that when a self-reference is introduced, a program might fall into an endless loop. To prove that this self-reference in P^{**}_T will never generate an endless loop, for any w , a mathematical induction proof will be built.

The simple “trick” to this demonstration is realizing that for any program of size N on $U_{P^{**}_T \circ P_T}$ to be well defined, it will merely depend on the programs in $U_{P^{**}_T \circ P_T}$ of size $< N$, or on P_T . Note that, as π_Ω is predetermined, it has a minimum size, thus there will be a maximum size $M_{(i)}$ for w such that if $|w| \leq M_{(i)}$, then $U(P^{**}_T \circ P_T \circ w)$ will never fall under case (i). With this, since P_T is given as total, then $P^{**}_T \circ P_T$ will always be well defined for sufficiently small w . To conclude the proof by induction, it is sufficient to note that if $P^{**}_T \circ P_T$ is well defined for every w of size $\leq k$, so $P^{**}_T \circ P_T$ will be well defined for every w of size $k + 1$. These are the key ideas of what will be done in theorem 10.1 below.

Once $P^{**}_T \circ P_T$ is total, $BB^+_{P^{**}_T \circ P_T}(N)$ and $\Omega_{P^{**}_T \circ P_T}$ can be defined, and therefore, as an immediate corollary of theorem 11.1, $BB^+_{P^{**}_T \circ P_T}(N)$ will be as uncomputable and as incompressible in relation to $U_{P^{**}_T \circ P_T}$ as $BB(N)$ is in relation to U — except for a constant. Note, thus, that $BB^+_{P^{**}_T \circ P_T}(N)$ also serve to “measure” the algorithmic complexity of programs on $U_{P^{**}_T \circ P_T}$: no output equal to or higher than $BB^+_{P^{**}_T \circ P_T}(N)$ can be achieved by any subprogram of size $\leq N$, but can always be calculated by a subprogram of size $\leq 2N + C$. See theorem 10.2.

The essential hypercomputational property that will thus be able to be simulated is that there are programs in $U_{P^{**}_T \circ P_T}$ that use finite lower approximations to $\Omega_{P^{**}_T \circ P_T}$ to occasionally calculate $BB^+_{P^{**}_T \circ P_T}$ values. At least regarding this aspect, it is as if the machine that computes function $BB^+_{P^{**}_T \circ P_T}(N)$ was a first order oracle Turing machine, only that now in relation to submachine $U_{P^{**}_T \circ P_T}$ and not in relation to U .

Theorem 10.1.

*Let P_T be a total function. Then, for every $w \in L$, $U(P^{**}_T \circ P_T \circ w)$ is a well-defined value, i.e. $U(P^{**}_T \circ P_T \circ w)$ is a total function of w ; or $U_{P^{**}_T \circ P_T}$ is a Turing submachine.*

PROOF.

a) *Case valid for every w where $|w| \leq k_0$:*

*As π_Ω is a known program, it has a size (in the language L of our choice). We also know program $P^{**}_T \circ P_T$, which will also have, therefore, a size. Hence,*

$$\exists k_0 (k_0 = \min \{ |\pi'_\Omega \circ P^{**}_T \circ P_T \circ w| ; w \in L \} - 1)$$

and

$$\forall w (|w| \leq k_0 \rightarrow w \neq \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho)$$

Through the definition of language L , determining whether

$$w = \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \vee w \neq \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$$

is a decidable problem. Thus, P^{**}_T can always recursively determine which option applies.

Thus,

$$\forall w (|w| \leq k_0 \rightarrow U(P^{**}_T \circ P_T \circ w) = U(P_T \circ w))$$

Since, by our initial assumption, P_T is a total function, then

$$\forall w \exists y (|w| \leq k_0 \rightarrow U(P^{**}_T \circ P_T \circ w) = y)$$

This concludes the first part of the mathematical induction.

b) Case valid for every w where $|w| \leq k$:

Our objective is to demonstrate that, in this case, it will also be valid for every w where $|w| = k + 1$.

First, take any arbitrary program w of size $k + 1$.

We have that either w is in form $\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$ or in any other form.

As in case a), knowing whether

$$w = \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \vee w \neq \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$$

is a decidable problem. Thus, P^{**}_T can always determine correctly any of these options. Below each possibility will be examined separately.

(i) In case $w = \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$:

It will suffice for P^{**}_T to calculate

$$U(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1))$$

and verify whether

$$U(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) \geq \rho$$

The crux of the matter is to prove that $U(P^{**}_T \circ P_T \circ w)$ needs to be defined for every w , wherein $|w| \leq |\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1$, so that $P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)$ always halts. By the inductive hypothesis, $U(P^{**}_T \circ P_T \circ w)$ is well defined for every w , with $|w| \leq k$. Therefore, since

$$|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| = k + 1$$

then $U(P^{**}_T \circ P_T \circ w)$ is defined for every w where

$$|w| \leq |\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1 = k$$

Thus, when P^{**}_T runs $P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)$, it will halt, never entering into a loop. Moreover, as ρ is given by w , determining whether or not

$$U \left(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1) \right) \geq \rho$$

will be a decidable problem. Thus, P^{**}_T can recursively verify whether

$$U \left(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1) \right) \geq \rho$$

In case $(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) \geq \rho$, by the definition of π'_{Ω} , we have that for every ρ' , if

$$U \left(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho'| - 1) \right) \geq \rho'$$

then $\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho'$ will always halt. Thus, if

$$U \left(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1) \right) \geq \rho$$

then $U(T \circ w) = U(T \circ \pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho)$ will always be well defined.

In case $(P_{\Sigma} \circ P^{**}_T \circ P_T \circ (|\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1)) < \rho$, we will proceed to case (ii).

Furthermore, due to the fact that P_T is total, one can always decide whether or not

$$U(T \circ w) > U(P_T \circ w)$$

Thus, $U_{P^{**}_T \circ P_T}(w)$ will always be well defined if w falls into case (i).

(ii) If case (i) does not apply to w :

It follows directly from the fact that P_T is total.

Thus, from (i) and (ii), for every w where $|w| = k+1$, $U(P^{**}_T \circ P_T \circ w)$ will be well defined.

c) Finishing the induction:

Through a) and b) it is concluded, by induction, that $U(P^{**}_T \circ P_T \circ w)$ is well defined for every $w \in L$.

Theorem 10.2.

Let N be a natural number. Let P_T be a total function. Then, for every N , there is a program in the form

$$\pi'_{\Omega} \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho$$

such that

$$U_{P^{**}_T \circ P_T} \left(\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \right) \geq BB^+_{P^{**}_T \circ P_T}(N)$$

and

$$\left| \pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \right| \leq 2N + C$$

where C is a constant.

THE KEY IDEA OF THE PROOF.

Let us use the fact that $\Omega_{P^{**}_T \circ P_T}$ is a number that contains the necessary information on programs on $U_{P^{**}_T \circ P_T}$ in order to compute function $BB^+_{P^{**}_T \circ P_T}(N)$. This is purposefully analogous to Ω and $BB(N)$. Furthermore, P^{**}_T was constructed to allow lower approximations to $\Omega_{P^{**}_T \circ P_T}$ to be used for $U_{P^{**}_T \circ P_T}$ to calculate $BB^+_{P^{**}_T \circ P_T}(N)$.

PROOF.

Through the previous theorem 10.1, it was proven that for any input w , $P^{**}_T \circ P_T \circ w$ halts. That is, $\forall w \exists y (U(P^{**}_T \circ P_T \circ w) = y)$. It follows then that function $U(P_\Sigma \circ P^{**}_T \circ P_T \circ N)$ will also be well defined for every N .

Let then N be an arbitrary natural number.

We will have that $U(P_\Sigma \circ P^{**}_T \circ P_T \circ N)$ give us a real binary number between 0 and 1, which is a finite lower approximation to $\Omega_{P^{**}_T \circ P_T}$.

Thus, from the definition of program π'_Ω ,

$$U \left(\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \right) = BB^+_{P^{**}_T \circ P_T}(N)$$

Also, clause (i) in the definition of program P^{**}_T assures us that if

$$U \left(P_\Sigma \circ P^{**}_T \circ P_T \circ (|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1) \right) \geq \rho$$

then

$$U_{P^{**}_T \circ P_T} \left(\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho \right) = U(\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho)$$

That is why we need to make $|\pi'_\Omega \circ P^{**}_T \circ P_T \circ 0^{|\rho|} 1 \circ \rho| - 1$ greater than or equal to N .

Let us, then, increase the size of $U(P_\Sigma \circ P^{**}_T \circ P_T \circ N)$, if necessary, through the following procedure: take the real number between 0 and 1 given by $U(P_\Sigma \circ P^{**}_T \circ P_T \circ N)$ and add (if necessary) enough trailing zeros until

$$\left| \pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \right| - 1 \geq N$$

We also have that, for every k, k' , if

$$k' \geq k$$

then

$$U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ k'}) \geq U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ k})$$

Therefore,

$$\begin{aligned} & U \left(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ 0} \middle| (\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ \right. \\ & \quad \left. \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| - 1 \right) \geq \\ & U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) = \\ & U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \end{aligned}$$

Thus, by the clause (i), as we have already remarked,

$$\begin{aligned} & U_{P^{**}_{T \circ P_T}} \left(\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \right) = \\ & U(\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0) \end{aligned}$$

But, since $U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0$ and $U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})$ are two equal real numbers, then

$$\begin{aligned} & U_{P^{**}_{T \circ P_T}} \left(\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \right) = \\ & U \left(\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0 \right) = \\ & U \left(\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) \middle| 1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) \right) = \\ & BB^+_{P^{**}_{T \circ P_T}}(N) \end{aligned}$$

Therefore, we obtain our intended program in the form $\pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \middle| 1 \circ \rho$ that calculates $BB^+_{P^{**}_{T \circ P_T}}(N)$ running on submachine $U_{P^{**}_{T \circ P_T}}$.

For the second part of the theorem, we have that, from the definition of language L and since $\pi'_{\Omega} \circ P^{**}_{T \circ P_T}$ has a finite size ≥ 1 , then, for every N ,

$$|U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})| \leq N + 1$$

So,

$$\left| \pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \Big|^{U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})} \Big|_{1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})} \right| \leq 2N + C$$

We will also have that, if

$$\left| \pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \Big|^{U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})} \Big|_{1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N})} \right| - 1 < N$$

and $0 \dots 0$ is minimal, then

$$\left| \pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \Big|^{U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0} \Big|_{1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0} \right| \leq N + 2$$

Thus, it leads us to the conclusion that there is a constant C such that

$$\left| \pi'_{\Omega} \circ P^{**}_{T \circ P_T \circ 0} \Big|^{U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0} \Big|_{1 \circ U(P_{\Sigma} \circ P^{**}_{T \circ P_T \circ N}) * 0 \dots 0} \right| \leq 2N + C$$

11. Sub-Uncomputability: recursive relative uncomputability

Now we will prove the crucial, yet simple, result that underpins this paper.

Let P'_T be a total function and $U_{P'_T}$ a Turing submachine. Then, we prove in theorem 11.1 that function $BB^+_{P'_T}(N)$ is relatively uncomputable by any program on submachine $U_{P'_T}$. Or: there is no subprogram that, for every input N , returns an output equal to $BB^+_{P'_T}(N)$. Once $BB^+_{P'_T}(N)$ is computable, then we say it is **sub-uncomputable**.

Note that, since $U_{P'_T}$ is a time-bounded Turing machine, $BB^+_{P'_T}$ promptly defines a time hierarchy [4, 25] in computational complexity theory. In fact, one can define BB^+ upon submachines that are not necessarily time-bounded Turing machines, so that, on the other hand, $BB^+_{U/f}(N)$ will be a function that defines arbitrary subrecursive hierarchies.

A more intuitive way to understand what is going on is to look for a program and concatenate its input, such as $U_{P'_T}(P * N)$ for instance. Where “ $*$ ” denotes the optimal functionalizing concatenation, and not necessarily the “concatenation” “ \circ ” defined. See section 2. In fact, this applies to any way to compress the information of P and N in an arbitrary subprogram. Anyhow, the theorem 11.1 will be proved for both “ \circ ” and for “ $*$ ”.

We avail ourselves of the same idea used in the demonstration of Chaitin’s incompleteness theorem [13]. Now, however, to demonstrate an uncomputability relative to the submachine $U_{P'_T}$.

When N is given as input to any program P , it comes in its compressed form with size $\cong H(N)$. In fact, we use the property

$$|P \circ N| \leq C + |P| + C' + \log_2 N + (1 + \epsilon) \log_2(\log_2 N)$$

Whereby,

$$|P \circ N| \cong C + H(N)$$

But, as already known by the AIT, for any constant C there is a big enough N_0 such that $C + H(N_0) < N_0$. Therefore, according to the definition of $BB^+_{P'_T}$, the output of $P \circ N_0$ when run on submachine $U_{P'_T}$, will be taken into account when one calculates $BB^+_{P'_T}(N_0)$. Thus, necessarily,

$$BB^+_{P'_T}(N_0) \geq U_{P'_T}(P \circ N_0) + 1 > U_{P'_T}(P \circ N_0)$$

Which will lead to contradiction, if P computes $BB^+_{P'_T}$ when running on submachine $U_{P'_T}$. The same holds for “*”.

Also, following the same argument, it can be shown promptly that $BB^+_{P'_T}$ is a **relatively incompressible**, or **sub-incompressible**, function by any sub-program smaller than or equal to N . That is, no program of size $\leq N$ running on $U_{P'_T}$ will result in an output larger than or equal to $BB^+_{P'_T}(N)$. This follows directly from the definition of the function $BB^+_{P'_T}$. Besides, if one proves sub-incompressibility of $BB^+_{P'_T}(N)$ at first hand, then sub-uncomputability follows as a corollary.

Note that nothing can be said as yet about the relative randomness (or incompressibility) of $\Omega_{P^{**} \circ P_T}$, which we aim indeed to study in further work.

Theorem 11.1.

Let N be an arbitrary natural number. Let P'_T be a total function. Let $U_{P'_T}$ be a Turing submachine. Then, function $BB^+_{P'_T}(N)$ is uncomputable by any program on $U_{P'_T}$. That is, one can prove a strict dominance:

$$\forall P \exists N_0 \forall N (N \geq N_0 \rightarrow U_{P'_T}(P * N) < BB^+_{P'_T}(N))$$

and

$$\forall P \exists N_0 \forall N (N \geq N_0 \rightarrow U_{P'_T}(P \circ N) < BB^+_{P'_T}(N))$$

PROOF.

Take an arbitrary program P in the language L . From the definition of language L and of U , we have that:

$$|P \circ N| \leq C + |P| + C' + \log_2 N + (1 + \epsilon) \log_2(\log_2 N)$$

*We have, from the definition of $P * N$ in 2, that $|P * N| \leq |P \circ N|$.*

Let $C + |P| + C' = C''$.

It is well-known that:

$$\exists N_0 \forall N (N \geq N_0 \rightarrow C'' + \log_2 N + (1 + \epsilon) \log_2(\log_2 N) < N)$$

Thus,

$$\exists N_0 \forall N (N \geq N_0 \rightarrow |P * N| \leq |P \circ N| < N)$$

Consequently, from the definition of $BB^+_{P'_T}$, we will have that there is N_0 such that

$$\forall N (N \geq N_0 \rightarrow U_{P'_T}(P \circ N) < U_{P'_T}(P \circ N) + 1 \leq BB^+_{P'_T}(N))$$

and

$$\forall N (N \geq N_0 \rightarrow U_{P'_T}(P * N) < U_{P'_T}(P \circ N) + 1 \leq BB^+_{P'_T}(N))$$

12. Conclusions

We have defined Turing submachines as another terminology for total Turing machines, but emphasizing the property that there are always non-reducibly more powerful Turing machines from which total Turing machines are subsystems [2].

Assuming P_T as program that computes a total function, a Turing submachine $U_{P^{**}_T \circ P_T}$ is built such that $U_{P^{**}_T \circ P_T}$ is a proper extension of U_{P_T} , by proving that $P^{**}_T \circ P_T$ is also a program that computes a total function. Furthermore, a relative and computable halting probability $\Omega_{P^{**}_T \circ P_T}$ — in the case, a time-bounded one — was defined regarding which was proved that there are finite lower approximations ρ that can be used by a program P when running on $U_{P^{**}_T \circ P_T}$ to compute values of $BB^+_{P^{**}_T \circ P_T}(N)$, so that $2N + C \geq |P \circ \rho| \geq N + 1$, where C is a constant. The submachine $U_{P^{**}_T \circ P_T}$ is a time-bounded Turing machine that increases the computation power of the time-bounded Turing machine limited by P_T . And it does that by recursively allowing some programs to calculate values of $BB^+_{P^{**}_T \circ P_T}(N)$ for lower approximations to $\Omega_{P^{**}_T \circ P_T}$ through a self-diagonalizing procedure.

We also proved that, for every Turing submachine $U_{P'_T}$, if P'_T is a program that computes a total function, then the computable function $BB^+_{P'_T}(N)$ is relatively uncomputable by any program running on $U_{P'_T}$ — in the “same manner” that the Busy Beaver function $BB'(N)$ is in relation to any program on U . Also, by the very definition of $BB^+_{P'_T}$, there cannot be any program of size $\leq N$ running on $U_{P'_T}$ that generates an output higher than or equal to $BB^+_{P'_T}(N)$. In other words, we have shown in the last theorem 11.1 that relative uncomputability and relative algorithmic incompressibility are ubiquitous phenomena for submachines. Thus, a more powerful containing system, from which a subsystem is a submachine, can promptly compute a problem of irreducible complexity compared to this subsystem. These results allow us to build a hierarchy of subrecursive classes built on relative uncomputability and relative incompressibility of the Busy Beaver function.

Moreover, since we have proved that $P^{**}_T \circ P_T$ is total, function $BB^+_{P^{**}_T \circ P_T}$ becomes sub-uncomputable by any program running on $U_{P^{**}_T \circ P_T}$, but still reachable by giving lower approximations to $\Omega_{P^{**}_T \circ P_T}$ as inputs to (at least) one kind of program running on $U_{P^{**}_T \circ P_T}$. Also, function $BB^+_{P^{**}_T \circ P_T}(N)$ is always incompressible by any program of size $\leq N$ running on $U_{P^{**}_T \circ P_T}$.

This was designed to mimic what a universal Turing machine can do with lower approximations to Ω with the purpose of calculating values of $BB'(N)$ while maintaining this function as a “yardstick” of irreducible algorithmic complexity/information — see [10–12].

Note that this does not mean that $\Omega_{P^{**}_T \circ P_T}$ is random/incompressible in relation to $U_{P^{**}_T \circ P_T}$, whereas it holds for relative halting probabilities across the Turing degrees for example [16]. In fact, not “all uncomputabilities” of a first order hypercomputer were relativized in relation to a universal Turing machine. We propose for further research to investigate to what extent one can make a Turing machine be made to “behave”, in relation to one of its submachines (or subsystems), as if it were a hypercomputer. Analogously, one can also ask if there is such a thing as “sub-randomness” of a bit string or a real number, in the same way that there is sub-uncomputability of a function. For example, these inquiries might encompass studying both the computational complexity of running such proper subrecursive extensions and its relations to Rice’s theorem and to Turing degrees within computability theory [20, 21].

Alongside with the pursuit for such limiting results, this kind of recursive relativization has already shown to be useful in order to build theoretical models of computable systems. For example, we have defined in [1] evolutionary models without the need of a hypercomputable environment (or “Nature”) that are fully analogous to Chaitin’s hypercomputable models of the open-ended cumulative (darwinian) evolution of software. That is, we have proved that the open-ended evolution of subprograms can be as fast as the open-ended evolution of programs, through simulating the uncomputability and incompressibility of the Busy Beaver function for Turing submachines.

References

- [1] F. S. Abrahão, *Metabiologia, Subcomputação e Hipercomputação: em direção a uma teoria geral de evolução de sistemas*, Ph.d. thesis, Universidade Federal do Rio de Janeiro (UFRJ), Brazil (2015).
- [2] F. S. Abrahão, The “paradox” of computability and a recursive relative version of the Busy Beaver function, in: C. Calude, M. Burgin (eds.), *Information and Complexity*, 1st ed., chap. 1, World Scientific Publishing, Singapore, 2016, pp. 3–15.
- [3] E. Allender, M. Koucký, D. Ronneburger, S. Roy, The pervasive reach of resource-bounded Kolmogorov complexity in computational complexity theory, *Journal of Computer and System Sciences* 77 (1) (2011) 14–40.
- [4] S. Arora, B. Barak, *Computational complexity: a modern approach*, 1st ed., Cambridge University Press, 2009.

- [5] S. K. Basu, On the structure of subrecursive degrees, *Journal of Computer and System Sciences* 4 (5) (1970) 452–464.
- [6] M. Burgin, *Measuring Power of Algorithms*, Computer Programs and Information Automata, Nova Science Publishers Inc, 2010.
- [7] C. S. Calude, *Information and Randomness: An algorithmmic perspective*, 2nd ed., Springer-Verlag, 2002.
- [8] C. S. Calude, G. J. Chaitin, WHAT IS... a Halting Probability ?, *Notices of the AMS* 57 (2) (2010) 236–237.
URL <http://www.ams.org/notices/201002/rtx100200236p.pdf>
- [9] W. Carnielli, F. A. Dória, Are the foundations of computer science logic-dependent?, in: C. Degremont, L. Keiff, H. Ruckert (eds.), *Dialogues, Logics and Other Strange Things—Essays in Honour of Shahid Rahman*, College Publications, London, 2008, pp. 87–107.
- [10] G. Chaitin, Life as Evolving Software, in: H. Zenil (ed.), *A Computable Universe: Understanding and Exploring Nature as Computation*, World Scientific Publishing, Singapore, 2012, pp. 277–302.
- [11] G. Chaitin, *Proving Darwin: making biology mathematical*, Vintage Books, 2013.
- [12] G. Chaitin, V. Chaitin, F. S. Abrahão, Metabiología: los orígenes de la creatividad biológica, *Investigación y Ciencia* (2014) 74–80.
- [13] G. J. Chaitin, Computational Complexity and Gödel’s incompleteness theorem, *SIGACT News* (9) (1971) 11–12.
- [14] G. J. Chaitin, A Theory of Program Size Formally Identical to Information Theory, *Journal of the ACM* 22 (3) (1975) 329–340.
- [15] B. J. Copeland, Hypercomputation, *Minds and Machines* 12 (4) (2002) 461–502.
- [16] R. Downey, D. R. Hirschfeldt, J. S. Miller, A. Nies, Relativizing Chaitin’s halting probability, *Journal of Mathematical Logic* 5 (2) (2005) 167–192.
- [17] H. B. Enderton, *A mathematical introduction to logic*, 2nd ed., Academic Press, 2001.
- [18] S. Hernández-Orozco, F. Hernández-Quiroz, H. Zenil, The Limits of Decidable States on Open-Ended Evolution and Emergence, in: *Fifteenth International Conference on the Synthesis and Simulation of Living Systems (ALIFE)*, Cancun, 2016, pp. 200–207.
URL <https://mitpress.mit.edu/books/proceedings-artificial-life-conference-2016>

- [19] S. Homer, A. Selman, *Computability and Complexity Theory*, 2nd ed., Springer, New York, 2011.
- [20] M. Hoyrup, The decidable properties of subrecursive functions, in: *LIPIcs-Leibniz International Proceedings in Informatics*, vol. 55, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016, pp. 1–13.
URL
<http://drops.dagstuhl.de/opus/volltexte/2016/6243/pdf/LIPIcs-ICALP-2016-108.pdf>
- [21] D. Kozen, Indexings of subrecursive classes, *Theoretical Computer Science* 11 (3) (1980) 277–301.
URL
<http://linkinghub.elsevier.com/retrieve/pii/0304397580900171>
- [22] D. Kozen, *Automata and Computability*, vol. 53, 1st ed., Springer-Verlag New York, 1997.
- [23] H. Lewis, C. H. Papadimitriou, *Elements of the Theory of Computation*, 2nd ed., Prentice-Hall, 1997.
- [24] M. Li, P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, 2nd ed., Springer Science & Business Media, New York, 1997.
- [25] C. H. Papadimitriou, *Computational Complexity*, Addison-Wesley Publishing, 1994.
- [26] K. W. Regan, Diagonalization, uniformity, and fixed-point theorems, *Information and Computation* 98 (1) (1992) 1–40.
- [27] H. E. Rose, *Subrecursion: Functions and Hierarchies*, Oxford University Press, 1984.
- [28] M. Sipser, *Introduction to the Theory of Computation*, vol. 2, Thomson Course Technology, Boston, 2006.
- [29] A. Syropoulos, *Hypercomputation: Computing beyond the Church-Turing Barrier*, 2008.
- [30] A. Turing, Systems of logic based on ordinals, *Proceedings of the London Mathematical Society*.
- [31] O. Watanabe (ed.), *Kolmogorov Complexity and Computational Complexity*, Springer-Verlag, 1992.
- [32] H. Zenil (ed.), *Randomness through Computation: Some Answers, More Questions*, World Scientific Publishing, 2011.